Does unemployment affect the number of Establishments? A Regional Analysis for U.S. States

Regional Studies, Vol. 36.4, pp. 389–398, 2002

Carfax Publishing
Taylor & Francis Group

Does Unemployment Affect the Number of Establishments? A Regional Analysis for US States

MARTIN A. CARREE

Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands and Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands. Email: m.carree@mw.unimaas.nl

(Received December 2000; in revised form December 2001)

CARREE M. A. (2002) Does unemployment affect the number of establishments? A regional analysis for US States, *Reg. Studies* **36**, 389–398. Empirical tests of the 'unemployment push' hypothesis that unemployed workers start up new enterprises to a greater extent than employed workers to escape unemployment have produced contradictory results. This is at least partially a consequence of studies neglecting the industrial organization of regions and not incorporating adequate variables to represent market opportunities. Using data of US establishments in retail and service industries with low entry barriers, we find little or no evidence for the 'unemployment push' hypothesis. US regions with relatively high unemployment rates do not show relatively strong subsequent increases in the number of establishments across a broad range of industries. Possible exceptions are used merchandise stores and automotive repair shops.

Consumer services Net entry Retailing Unemployment

CARREE M. A. (2002) Le chômage, est-ce qu'il influe sur le nombre d'établissements? – une analyse régionale pour les Etats-Unis, *Reg. Studies* **36**, 389–398. Des tests empiriques de l'hypothèse qui affirme que la création d'entreprise s'explique plutôt par les chômeurs que par les actifs occupés afin de permettre à ces gens-là de contourner le chômage, ont réalisé des réultats contradictoires. Cela s'explique du moins en partie par le fait que des études n'ont pas considéré la structure industrielle des régions et n'ont pas incorporé des variables suffisantes pour représenter les possibilités offertes par le marché. A partir des données auprès des établissments dans le commerce et les services dont les barrières à l'entrée sont très basses, il s'avère presque pas de preuves à l'appui de cet hypothèse. Les régions aux Etats-Unis dont le taux de chômage est relativement élevé ne montrent pas d'augmentations ultérieures relativement fortes du nombre d'établissements à travers des industries diverses. A quelques exceptions près, à savoir les vendeurs d'objets d'occasion et les ateliers de réparation automobile.

Services aux particuliers Entrée nette
Commerce de détail Chômage

CARREE M. A. (2002) Beeinflußt Erwerbslosigkeit die Anzahl kommerzieller Unternehmen? Eine Regionalanalyse für die Staaten der USA, *Reg. Studies* **36**, 389–398. Empirische Untersuchungen der Hypothese des Anstosses, den Erwerbslosigkeit gibt, neue Unternehmen aufzuziehen, indem sie eher erwerbslose, Verdienst suchende Arbeiter dazu veranlaßt, als solche, die erwerbstätig sind, haben widersprüchliche Ergebnisse gezeitigt. Die ist zumindest teilweise Studien zuzuschreiben, welche die industrielle Organisation von Regionen vernachlässigen und versäumen, angemessene Variable zur Darstellung von Marktchancen einzuschließen. Der Autor benutzt Daten amerikanischer Unternehmen des Einzelhandels und der Dienstleistungindustrie mit mäßigen Einstiegsanforderungen, findet jedoch wenig oder keine Beweise für die Hypothese, daß Erwerbslosigkeit einen Anstoß darstellt. Regionen der USA mit verhältnismäßig hohen Erwerbsloigkeitsraten zeigen bei breiter Palette von Industrien später durchweg keine unverhältnismäßig starke Zunahme der Anzahl von Betriebsneugründungen auf. Gebrauchtwarenläden und Werkstätten für Kraftfahrzeugreparaturen stellen mögliche Ausnahmen dar.

Verbraucherdienste Netzeintrag Einzelhandel
Erwerbslosigkeit

INTRODUCTION

There is considerable interest in and controversy about the relationship between spatial variation in unemployment and that in the (net) entry of firms (e.g. REYNOLDS *et al.*, 1994). It has been claimed that higher rates of unemployment lead to subsequent higher rates of firm entry, *ceteris paribus.* The argument is that unemployed workers seek self-employment as an escape out of unemployment. This 'unemployment push' argument can also be described as a 'desperation' argument: even after correcting for the attractiveness

0034-3404 print/1360-0591 online/02/040389-10 DOI: 10.1080/00343400220131151
http://www.regional-studies-assoc.ac.uk

of becoming self-employed, unemployment still will have a positive effect on subsequent entry of (small) enterprises or establishments and/or on subsequent self-employment rates. The empirical evidence for the 'unemployment push' hypothesis has, thus far, been inconclusive. STOREY, 1991, finds that, in general, time-series analyses point to unemployment being positively related to indices of new firm formation, whereas cross-sectional studies appear to indicate the opposite (see also FOTI and VIVARELLI, 1994, p. 83).[1]

Recent cross-sectional, i.e. regional, studies that have found that high levels of unemployment tend to depress firm births include AUDRETSCH and FRITSCH, 1994, for 75 geographical regions in Germany, and REYNOLDS *et al.*, 1995, for 382 US labour markets. These results indicate that a relatively depressed business environment is found in areas of unemployment. However, AUDRETSCH and FRITSCH, 1999, show, for Germany, that this negative effect disappears when the industrial organization of firm entries is taken into account. They find that, out of 31 manufacturing industries, none shows a significant negative effect of unemployment on the new-firm start-up rate while 12 show a significant *positive* effect. One would expect such positive effects to be found for small-scale, easy-to-enter industries and their results give some support for that, with, for example, the 'leather industry' having the highest impact of the unemployment rate on the start-up activity.

The result found by AUDRETSCH and FRITSCH, 1999, that there is a systematic tendency for the impact of regional factors (and specifically unemployment) on new firm formation to differ across industries has important consequences for research into the validity of the 'unemployment push' hypothesis. A first consequence is that the negative effect found in regional analyses like REYNOLDS *et al.*, 1995, for the US may be the result of ignoring the role of the industrial organization of the regional economies. A second consequence is that we should expect industries which are relatively easy to enter, like retailing and service industries, to show a strong positive effect of unemployment.[2] In this study we seek to find evidence for these two hypotheses using a data set of 'single line' retail and service industries in US states. The approach takes into account incentives for firms to enter (and not to exit) the specific industries to avoid having to use the unemployment ratio as an index of the business environment. The result of the analysis is that, largely, *no effect* of unemployment on the net entry rate of establishments in US retail and service industries is found.

The rest of this paper is organized as follows. In the next section we introduce our model explaining regional variation in the change in the number of establishments in retail and consumer service industries. In the third section we discuss the data set of 16 industries in US states. It is followed by the estimation results in the fourth section. Finally, conclusions will be drawn.

THE MODEL

An important element that retail and consumer service industries have in common is that they serve *local* markets. This places an important upper limit on the size of establishments because consumers are not willing to travel large distances. The internet and TV home shopping may have changed this notion to some extent as large companies like Amazon are not bound by geographical markets. However, the data set under consideration ends in 1997 and therefore largely precedes the Internet and, in addition, only a few retail activities are suitable for TV home shopping. The model applied consists of two parts. In the first part, the number of establishments in a state is predicted using data on population, income, urbanization and age distribution. The residual of this regression is considered as an indicator of the opportunities present for new enterprises to enter. In the second part, the change in the number of establishments is explained from changes in population and income, from the residual of the first regression equation and from the level of unemployment.

Not all establishments are also firms. A firm is defined as an aggregation of all establishments owned by a parent company and it may consist of a single, independent establishment or it can include subsidiaries or other branch establishments under the same ownership and control. However, the retail and service industries that we concentrate upon have low ratios of the number of establishments to firms indicating that the barriers to unemployed to start up a venture are not severely heightened as a results of the dominance of few large companies (franchisers).

The key determinants of the number of establishments that a region can 'carry' are the population, their income, their distribution in space and their distribution by age.[3] We consider the following variables to predict the number of establishments in a certain state in 1993 (N). The first is *Pop*, the total population of a state (in thousands). The second and third are variables indicating the purchasing power of that population: they are *Rinc*, the real income per capita (in dollars); and *Pov*, the percentage of people who are 'in poverty' according to the Bureau of Census definition. The fourth and fifth are variables indicating the extent to which people are geographically clustered, they are *Urban*, the percentage of people living in 'urban areas' and *Dens*, the population density (total population divided by the total state area). The last variable concerns the age distribution of the population: *Over*65, the percentage of population who are 65 years or older. All the variables are measured in the year 1993.[4] No unemployment variable is incorporated at this stage. The first equation is:

$$N_{ij} = \alpha_i Pop_j^{\beta_i} Rinc_j^{\gamma_i} Dens_j^{\delta_i} \exp(\zeta_i Pov_j + \eta_i Urban_j + \theta_i Over65_j + \varepsilon_{ij}) \quad (1)$$

with i and j being the indices for industry and state. Note that we include the three variables expressed in percentages in a different (exponential) form than the other three variables. The error term ε_{ij} is assumed to be independently and identically distributed across industries and regions.[5] We estimate equation (1) by first taking logarithms:

$$\ln(N_{ij}) = c_i + \beta_i LnPop_j + \gamma_i LnRinc_j + \delta_i LnDens_j + \zeta_i Pov_j + \eta_i Urban_j + \theta_i Over65_j + \varepsilon_{ij} \quad (2)$$

with $c_i = \ln(\alpha_i)$ being the industry-specific constant. We expect the 'elasticity' of population, β_i, to be close to unity. An increase in population usually implies that more local markets are available for entry. We expect that a high level of (real) income per capita is associated with a high demand ($\gamma_i > 0$) while a high level of people being 'in poverty' leads to less demand ($\zeta_i < 0$). For industries selling luxury products, the first variable will probably be more important, while for industries selling products satisfying 'basic needs' the second variable may be more relevant. We expect that in regions with a higher level of concentration of the population there will be fewer establishments necessary because the average distance between consumers and establishments is lower. Therefore, we expect that, in general, population density and urbanization both have a negative effect on the number of establishments ($\delta_i < 0$ and $\eta_i < 0$).[6] The effect of the last variable, the percentage of people with age 65 or higher, on the number of establishments may be positive or negative depending upon the industry.

The residuals of equation (2), $\hat{\varepsilon}_{ij}$, can be considered as indicators of the (lack of) 'room' for new establishments in a region. A positive value for $\hat{\varepsilon}_{ij}$ indicates that there are more establishments in a region than would have been expected on the basis of the population, income, density and age distribution variables. It implies that there is likely to be less room for additional establishments when compared to regions in which $\hat{\varepsilon}_{ij}$ is negative. We use the residual as one of the explanatory variables in the change of the number of establishments in the 1993–97 period.[7] Its effect should be negative in case of 'error correction' (or regression-towards-the-mean). Additional explanatory variables are the change in population (*DLnPop*), the change in real income (*DLnRinc*) and the change in the poverty rate (*DPov*) in the same period. Changes in the degree of urbanization and the age distribution are not taken into account as these barely change over time. The last variable we take into account is the regional unemployment rate (*Unemp*), being the percentage of workers who report themselves as being unemployed. According to the 'unemployment push hypothesis' this variable should have a positive effect on the change in the number of establishments because: (1) people choose self-employment over unemployment; and (2) incumbent entrepreneurs are less likely to find alternative occupations when the unemployment rate is high.[8] The equation we estimate is:

$$\Delta\ln(N_{ij}) = \kappa_i + \lambda_i DLnPop_j + \mu_i DLnRinc_j + \nu_i DPov_j + \pi_i \hat{\varepsilon}_{ij-1} + \rho_i Unemp_{j-1} + \nu_{ij} \quad (3)$$

with ν_{ij} being the error term. The estimated residual $\hat{\varepsilon}$ and the unemployment rate are measured in 1993 and therefore have an additional '-1' subscript. A process of 'error correction' would imply that $\pi_i < 0$. The 'unemployment push' hypothesis for an industry i is that ρ_i exceeds zero. The values of λ_i and μ_i are expected to be positive because increases in population and incomes generally lead to increased market room. The value of ν_i is expected to be negative because an increase in the poverty rate will generally lead to a decrease of market room.

An alternative to a two-step procedure, in which the residuals of equation (2) are incorporated into equation (3), is a one-step procedure.[9] By substituting ε_{ij} from equation (2) into equation (3) we get:

$$\begin{aligned}\Delta\ln(N_{ij}) = {} & \kappa_i - \pi_i c_i + \lambda_i DLnPop_j + \mu_i DLnRinc_j \\ & + \nu_i DPov_j + \rho_i Unemp_{j-1} + \pi_i \ln(N_{ij-1}) \\ & - \pi_i \beta_i LnPop_{j-1} - \pi_i \gamma_i LnRinc_{j-1} \\ & - \pi_i \delta_i LnDens_{j-1} - \pi_i \zeta_i Pov_{j-1} - \pi_i \eta_i Urban_{j-1} \\ & - \pi_i \theta_i Over65_{j-1} + \nu_{ij}\end{aligned} \quad (4)$$

The advantage of using the one-step procedure is that the interaction between short term effects (those of equation (3)) and long term effects (those of equation (2)) can be taken into account. One disadvantage of the procedure is that the extent to which the long term effects can be accurately estimated depends upon the estimate of the speed of error-correction π. Another disadvantage is that no less than 12 parameters have to be estimated using data for the US states. An important difference between equations (3) and (4) is that the error-correction parameter π is connected to the residual $\hat{\varepsilon}$ in the first equation and to $\Delta\ln(N_{ij-1})$ in the latter. It is uncertain which of the two estimation procedures is the preferred one in the current study and, therefore, we show estimation results in the fourth section both for the one-step and two-step procedure.

DATA AND ESTIMATION ISSUES

We consider 16 'limited product line' industries in retailing and consumer services.[10] The industries are presented in Table 1 together with the number of establishments having payroll in the US in 1997, the average number of employees in that same year, the percentage increase in the number of establishments and the ratio of the number of establishments over the number of (employer) firms. We use data for each of these industries for 49 states (excluding Washington

Table 1. Summary statistics for retail and consumer services industries, 1997

SIC	Description	Number of establishments	Average no. employees	Change 1993–97 (%)	Establishment/ firms
5540	Gasoline service stations	95,847	7·5	−5·5	1·74
5570	Motorcycle dealers	3,608	8·0	0·2	1·03
5736	Musical instrument stores	4,501	6·2	8·7	1·17
5930	Used merchandise stores	23,780	5·0	14·3	1·24
5941	Sporting goods and bicycle shops	24,527	7·2	2·4	1·21
5942	Book stores	12,301	9·7	−8·9	1·80
5944	Jewellery stores	29,459	5·3	6·8	1·50
5992	Florists	26,505	4·7	−2·6	1·05
5995	Optical goods stores	15,224	4·9	10·3	2·18
7230	Beauty shops	82,228	4·9	−1·2	1·13
7530	Automotive repair shops	144,068	4·4	9·5	1·06
7620	Electrical repair shops	19,846	7·2	−6·0	1·12
7830	Motion picture theatres	6,317	19·1	−6·7	2·66
7840	Video tape rental	23,182	6·4	5·5	1·69
7910	Dance studios, schools and halls	5,743	5·0	11·9	1·02
7930	Bowling centres	5,526	15·7	−8·0	1·14

Notes: The industry classification is that of SIC 1987. The third column presents the total number of establishments having payroll in the US in 1997. The fourth column presents the average number of employees per establishment having payroll in the US in 1997. The fifth column presents the percentage change in the number of establishments from 1993 to 1997. The last column presents the ratio of the number of establishments to the number of (employer) firms. None of the 16 industries had more than 10 establishments with 250 or more employees in the US in 1997.

Source: *County Business Patterns 1997*, US Census Bureau, US Small Business Administration.

DC and the state of Alaska). We selected the industries from those available from the US Census Bureau County Business Patterns which had the following characteristics in common. First, the size of establishments should generally be small. No industries were included where there were more than 10 establishments in the US with more than 250 employees. Second, the industries should be recognizable as being 'limited-lined'. Third, there should be at least 3,500 establishments in the US in total (in 1997). Fourth, the industries should not be food retailing to avoid the effects of competition with grocery stores. Fifth, the industries should have at least half of the establishments run by firms with less than 500 employees. We require this in order to exclude industries dominated by a few firms.

The average number of employees in most of the industries was below 10 (in 1997) with the exception of 'motion picture theatres' and 'bowling centres'. These two consumer service industries are not examples of low barriers to entry industries but were nevertheless included as they fulfilled the requirements and could serve as a check. The average number of establishments per firm was usually low, with the exception of motion picture theatres (2·66) and optical goods stores (2·18). For industries like motorcycle dealers, florists, automotive repair shops and dance studios, schools and halls, firms almost always consist of one establishment only.[11] In Table 2 we show descriptive statistics of the explanatory variables of equation (2) in 1993.

Table 3 shows the distribution of some key variables across the states. In the second to fourth columns of the table, the total number of establishments (for the 16 industries) per 1,000 inhabitants, the poverty rate and the degree of urbanization are reported for the year 1993. The three states with the lowest ratio of establishments to population are California, District of Columbia and Louisiana. For the state of Cailifornia this may be caused by the high degree of urbanization (92·6%) and for the state of Louisiana it may be a result of the high poverty rate (25·5%). The District of Columbia has both a high degree of urbanization (100%) and a high poverty rate (22·7%). These appear to be two important factors explaining the very low number of establishments per capita in the US capital. The three states with the highest ratio of establishments to population are South Dakota, Vermont and Wyoming. These states are characterized by relatively low poverty rates

Table 2. Descriptive statistics of explanatory variables

Variable	Description	Mean	Std. dev.	Min.	Max.
LnPop	Logarithm of total population	8·10	1·00	6·15	10·35
LnRinc	Logarithm of real income per capita	9·44	0·13	9·18	9·77
Poverty	Percentage of people 'in poverty'	13·85	3·92	8·8	25·5
Urban	Percentage of people in urban areas	68·21	14·85	32·2	92·6
LnDens	Logarithm of population density	−2·57	1·27	−5·34	−0·04
*Over*65	Percentage of people with age 65+	12·86	1·80	9·0	19·0
Unemp	Percentage of workers unemployed	6·30	1·49	2·7	10·9

Note: All variables are measured in 1993.

Source: US Census Bureau, US Bureau of Labor Statistics.

Table 3. Summary statistics for US states

State	Establishments per capita	Poverty	Urban	Change: establishments per capita	Unemployment
Alabama	2·01	17·0	60·4	−0·02	7·6
Alaska	2·06	9·8	67·5	0·13	7·7
Arizona	1·92	15·7	87·5	−0·13	6·3
Arkansas	2·05	17·6	53·5	−0·02	6·2
California	1·66	17·5	92·6	−0·05	9·4
Colorado	2·35	9·9	82·4	−0·00	5·3
Connecticut	2·32	9·7	79·1	−0·07	6·3
Delaware	2·27	8·8	73·0	−0·09	5·3
District of Columbia	1·40	22·7	100·0	0·05	8·6
Florida	2·31	16·1	84·8	−0·13	7·0
Georgia	1·98	15·1	63·2	−0·06	5·8
Hawaii	1·94	9·3	89·0	−0·06	4·3
Idaho	2·37	13·4	57·4	−0·17	6·2
Illinois	1·82	13·9	84·6	−0·02	7·5
Indiana	2·03	12·6	64·9	−0·02	5·4
Iowa	2.51	10·9	60·6	−0·05	4·0
Kansas	2·33	13·0	69·1	−0·08	5·0
Kentucky	1·91	19·5	51·8	−0·04	6·2
Louisiana	1·66	25·5	68·1	0·09	7·5
Maine	2·24	12·8	44·6	0·07	7·9
Maryland	1·90	10·7	81·3	−0·04	6·2
Massachusetts	2·16	10·2	84·3	0·01	6·9
Michigan	1·87	14·4	70·5	0·03	7·1
Minnesota	2·18	12·1	69·9	0·02	5·1
Mississippi	1·82	23·1	47·1	−0·00	6·4
Missouri	2·27	15·8	68·7	−0·07	6·5
Montana	2·64	13·4	52·5	0·01	6·1
Nebraska	2·62	9·9	66·1	−0·09	2·7
Nevada	1·77	11·8	88·3	−0·11	7·3
New Hampshire	2·59	8·8	51·0	0·02	6·6
New Jersey	2·14	10·1	89·4	0·01	7·5
New Mexico	2·04	20·0	73·0	−0·05	7·7
New York	1·79	16·4	84·3	0·02	7·8
North Carolina	2·03	14·8	50·4	−0·02	4·9
North Dakota	2·60	11·2	53·3	−0·02	4·4
Ohio	1·92	13·2	74·1	−0·04	6·5
Oklahoma	1·87	18·4	67·7	−0·02	6·1
Oregon	2·04	11·7	70·5	−0·08	7·3
Pennsylvania	2·01	12·5	68·9	0·02	7·1
Rhode Island	2·24	11·3	86·0	0·07	7·8
South Carolina	2·01	17·1	54·6	−0·01	7·6
South Dakota	2·75	14·6	50·0	−0·01	3·6
Tennessee	1·96	17·1	60·9	−0·05	5·7
Texas	1·92	18·3	80·3	−0·09	7·2
Utah	1·82	9·4	87·8	−0·03	3·9
Vermont	2·88	9·4	32·2	0·00	5·5
Virginia	2·08	10·0	69·4	−0·03	5·1
Washington	2·02	11·7	76·4	−0·06	7·6
West Virginia	1·91	21·1	36·1	0·01	10·9
Wisconsin	2·23	10·9	65·7	−0·04	4·7
Wyoming	3·12	11·0	65·0	0·08	5·5

Note: The number of establishments per capita is the total number of establishments having payroll for the 16 industries per 1,000 inhabitants. All variables are measured in 1993, except for the change in the number of establishments per capita (1993–97).

Source: See Table 2.

(especially Vermont and Wyoming) and low urbanization (especially South Dakota and Vermont).

The last two columns of Table 3 show the change in the total number of establishments (for the 16 industries) over the 1993–97 period and the unemployment rate in 1993. The correlation coefficient between these two variables across the states equals 0·26 (0·19 if Alaska and Washington DC are left out). This correlation appears to be caused by the states with low unemployment rates having declining numbers of establishments per capita (all states with 5·0% or less unemployment in 1993 have a decrease in these numbers over the 1993–97 period) rather than states with high unemployment rates having relatively strong increases in the number of establishments per capita. In fact, the two states with highest unemployment rates in 1993 (California and West Virginia) show no important change in the number of establishments per capita. The positive bivariate correlation between unemployment and subsequent change in the number of establishments per capita might be considered to indicate support for the 'unemployment push' hypothesis. However, important additional explanatory variables for this change, like population growth, change of disposable incomes and 'market room' (error correction), have to be taken into account as well to arrive at an adequate test.

We will show estimation results using the Stepwise regression technique (SPSS for Windows 10.0.5) for the two-step procedure and require variables to be significant at the 5% level to be included. The reason for using the Stepwise regression technique is that the variables of *LnRinc* and *Pov* and the variables of *LnDens* and *Urban* are strongly correlated. Including both could result in imprecise estimates. An additional advantage is that the number of coefficients to report is reduced. A potential problem of Stepwise regression is that no one variable is significant at entry but that a combination of two is. We checked for this by initially using a much higher percentage for the significance level. Only once did such a problem occur (industry 762 for equation (3)). For the one-step procedure the Stepwise regression technique is less suitable because equation (4) is non-linear in the parameters.

ESTIMATION RESULTS

We will first discuss the results of the two-step procedure. The Stepwise least squares results of equation (2) for each of the industries are shown in Table 4. We find the explanatory power of the regression in terms of R^2 to be 0·93 or higher with the exception of bowling centres (SIC 793). For each of the industries, the population has a significant effect (scale effect): the value of the estimated coefficient is usually close to unity. The effect of real income is significant in only five out of 16 industries. Equally strong effects are found for motorcycle dealers (SIC 557), jewellery stores (SIC

Table 4. Determinants of the number of establishments in 1993

Variable	554	557	5736	593	5941	5942	5944	5992	5995	723	753	762	783	784	791	793
Constant	0·36	−11·59	−4·05	−2·85	−1·08	−2·05	−7·89	−4·67	−3·31	−9·14	−0·79	−2·50	−3·46	−8·32	−3·40	−5·24
	(1·7)	(3·0)	(14·9)	(6·8)	(3·0)	(8·6)	(3·7)	(3·3)	(13·7)	(2·9)	(3·7)	(12·8)	(9·9)	(3·6)	(7·0)	(6·1)
LnPop	0·933	0·973	1·011	1·001	0·913	0·944	0·958	0·954	1·093	0·969	1·000	0·999	0·975	0·941	0·973	1·034
	(32·0)	(25·0)	(32·2)	(22·4)	(22·1)	(31·5)	(32·3)	(54·5)	(35·8)	(25·7)	(51·8)	(41·7)	(24·1)	(32·8)	(19·0)	(13·2)
LnRinc		0·850					0·635	0·319		0·824				0·817		
		(2·1)					(2·8)	(2·0)		(2·4)				(3·2)		
Poverty		−0.019	−0·028		−0·053	−0·030			−0·031	−0·035	−0·016		−0·030			−0·080
		(1·9)	(4·2)		(6·0)	(3·9)			(4·0)	(3·1)	(4·0)		(3·5)			(4·8)
Urban	−0·0095	−0·0092						−0·0078						−0.0119		
	(4·9)	(3·7)						(5·3)						(5·0)		
LnDens		−0·212	−0·106	−0·144	−0·095						−0·058		−0·208		0·136	−0·294
		(6·7)	(4·4)	(4·1)	(3·0)						(3·9)		(6·8)		(3·4)	(4·9)
Over65								0·030		0·074	0·018			−0·038		0·126
								(3·4)		(4·5)	(2·1)			(2·6)		(3·7)
R^2	0·961	0·955	0·970	0·929	0·937	0·957	0·963	0·987	0·966	0·963	0·990	0·974	0·941	0·966	0·933	0·818

Note: The dependent variable is the logarithm of the number of establishments in 1993. *T*-values between brackets. The state of Alaska and Washington DC are left out, leaving 49 observations for each regression. The results have been computed using Stepwise regression with a 5% significance level at entry and a 10% significance level at exit.

Table 5. Determinants of the change in the number of establishments, 1993–97

Variable	554	557	5736	593	5941	5942	5944	5992	5995	723	753	762	783	784	791	793
Constant	−0.047	−0·071	0·052	0·083	−0·034	−0·144	0·021	0·088	0·098	−0·032	0·074	−0·043	−0·050	−0·030	0·175	−0·081
	(8·4)	(2·7)	(2·1)	(3·4)	(1·8)	(6·0)	(1·5)	(2·2)	(4·8)	(3·4)	(6·6)	(1·9)	(4·3)	(1·4)	(9·7)	(7·7)
DlnPop		1·189	0·974		0·524	1·389	1·120		0·734			0·913		1·235		
		(2·7)	(2·1)		(2·1)	(3·5)	(4·3)		(2·2)			(2·9)		(3·3)		
DlnRinc				1·097	0·894						0·626	−0·899				
				(2·6)	(2·8)						(3·2)	(2·4)				
Dpoverty		−0·023				−0·019			0·014					−0·015	0·023	
		(3·1)				(2·9)			(2·4)					(2·4)	(3·2)	
Error93	−0·131		−0·401					−0·319	−0·154			−0·188	−0·213	−0·373	−0·241	
	(4·1)		(3·8)					(3·5)	(2·5)			(2·7)	(3·7)	(4·6)	(4·4)	
Unemp								−0·015								
								(2·3)								
R^2	0·264	0·248	0·325	0·126	0·252	0·288	0·285	0·256	0·275	0·000	0·179	0·258	0·225	0·406	0·406	0·000

Notes: The dependent variable is the number of establishments from 1993 to 1997. *T*-values between brackets. The state of Alaska and Washington DC are left out, leaving 49 observations for each regression. The results have been computed using Stepwise regression with a 5% significance level at entry and a 10% significance level at exit. '*Error*93' is the residual ε_{ij} from equation (2).

5944), beauty shops (SIC 723) and video tape rental (SIC 784). The first three are indeed industries which sell 'luxuries'. The effect of poverty is significantly negative in nine out of 16 industries. The strongest effects are found for sporting goods and bicycle shops (SIC 5941) and bowling centres (SIC 793). Note that, in four industries, neither of the two income variables has a significant impact. These are gasoline service stations (SIC 554), used merchandise stores (SIC 593), electrical repair shops (SIC 762) and dance studios, schools and halls (SIC 791). The effect of the urbanization variable is significant in only four out of 16 industries while the population density variable is significant in eight cases, but once with the 'wrong sign'. This one industry is dance studios, schools and halls (SIC 791).[12] Finally, the percentage of elderly people (age 65 or older) has a significant effect in five industries. A strong positive effect is found for beauty shops (SIC 723) and bowling centres (SIC 793). A negative effect is found for video tape rental (SIC 784). Overall, the results seem to be in line with expectations. The residuals of the regressions are now used in the second stage.

The estimation results for equation (3) can be found in Table 5. We find that the change in population has a positive and significant effect on the change in the number of establishments in eight cases. For only six industries do we find either the change in real income to have a positive effect or the change in the 'poverty' percentage to have a negative effect. The 'error correction' effect of the residual of equation (2) is negative and significant in eight out of 16 industries. The speed of 'error correction' is especially fast for musical instrument stores (SIC 5736), florists (SIC 5992) and video tape rental (SIC 784). The results do show that there are region-specific differences in many industries which appear not to diminish in the time period under consideration. It may be due to products sold in certain industries being more popular in some parts of the country than in others. The most important result in Table 5 for this study concerns the effect of the unemployment rate: there appears to be *none*.[13] There is one industry (florists) in which there is a significant effect but it has a negative sign and with 16 industries it is not at all unlikely to find one industry to have a significant effect at the 5% significance level.[14]

Although the results on the 'push hypothesis' for easy-to-enter industries leave little doubt, we pool the industries to consider an 'overall effect'. This pooling is not warranted by the heterogeneity we find in Table 4 and 5 but it may give some additional insights. We show the estimation results in Table 6. When considering all 16 industries we find a negative but insignificant effect of unemployment on the subsequent net entry rate of establishments. When we exclude those industries with less than 10,000 establishments (in 1997), being motorcycle dealers (SIC 557), musical instrument stores (SIC 593), motion picture theatres (SIC 783), dance studios, schools and halls (SIC 791) and bowling centres (SIC 793) we find that the coefficient rises somewhat but that the effect of the unemployment rate is far from significant.

Now we turn to the one-step procedure. The least squares results of equation (4) are presented in Table 7. We will focus on the error-correction parameter π_i and the effect of unemployment ρ_i. The results for the degree of error-correction are completely in line with those in the two-step procedure. The industries for which the error-correction coefficient is significant at the 5% significance level are the same in the two procedures. The results of the effect of unemployment are somewhat different than those found in the two-step procedure. In 13 out of 16 industries the effect is insignificant at the 10% significance level; in one industry it is *positive* and significant at the 10% level but not at the 5% level (used merchandise stores, SIC 593); in one industry it is *positive* and significant at the 5% level (automotive repair shops, SIC 753); and in one industry it is *negative* and significant at the 5% level (bowling centres, SIC 793). Both used merchandise stores and automotive repair shops can be considered industries with very low barriers to enter. The average number of employees per establishment for these two industries are just 5·0 and 4·4, respectively (see Table 1). The industry of bowling centres has larger-scaled establishments with an average number of employees of 15·7 (see Table 1). Starting a firm in this industry is more demanding for the potential entrepreneur. The results of the one-step procedure confirm that of the two-step procedure – that the large majority of low barrier industries do not show an effect of unemployment. However, there is a qualification in that the two-step procedure does provide empirical support for the 'unemployment push' hypothesis in one or two very low barrier industries.

Table 6. Determinants of the change in the number of establishments, aggregated

	DlnPop	*DlnRinc*	*Dpovert*	*Error93*	*Unemp*	R^2	*N*
All industries	0·657	0·021	−0·002	−0·101	−0·003	0·460	784
	(7·1)	(0·2)	(1·3)	(6·6)	(1·2)		
Industries with 10,000 or more establishments	0·634	0·201	−0·002	−0·072	−0·000	0·480	539
	(6·5)	(1·5)	(1·1)	(3·9)	(0·2)		

Notes: *T*-values between brackets. *N* is the number of observations, equal to 49 states times 16 industries for the 'all industries case'. '*Error*93' is the residual $\hat{\varepsilon}_{ij}$ from equation (2).

Table 7. Estimation results for the one-step procedure (equation (4))

Variable	554	557	5736	593	5941	5942	5944	5992	5995	723	753	762	783	784	791	793
Constant	−1·35	3·40	1·40	2·14	1·10	2·79	−0·18	1·11	−1·73	0·44	−1·16	0·25	−0·25	0·88	−3·00	−0·37
	(1·8)	(1·3)	(0·5)	(1·1)	(0·8)	(1·3)	(0·1)	(0·8)	(0·8)	(0·3)	(1·9)	(0·1)	(0·1)	(0·5)	(1·3)	(0·2)
DlnPop	0·010	1·803	1·456	0·210	0·599	1·009	0·977	−0·199	0·665	0·013	0·045	0·770	0·402	0·841	−0·141	−0·136
	(0·0)	(3·1)	(2·3)	(0·5)	(1·8)	(2·1)	(2·7)	(0·7)	(1·4)	(0·0)	(0·3)	(1·8)	(0·9)	(1·9)	(0·3)	(0·3)
DlnRinc	0·283	−0·602	−0·685	1·182	1·090	−0·393	0·206	0·417	0·167	0·242	0·722	−0·613	−0·118	−0·137	0·161	0·187
	(1·3)	(0·9)	(0·9)	(2·4)	(2·7)	(0·7)	(0·5)	(1·2)	(0·3)	(0·7)	(4·3)	(1·3)	(0·3)	(0·3)	(0·2)	(0·4)
Dpovert	0·001	−0·011	−0·008	−0·002	−0·001	−0·017	−0·013	−0·008	0·012	−0·004	0·001	0·002	−0·005	−0·008	0·008	0·004
	(0·4)	(1·2)	(0·7)	(0·3)	(0·1)	(1·9)	(2·0)	(1·5)	(1·4)	(0·7)	(0·3)	(0·2)	(0·7)	(1·1)	(0·9)	(0·6)
Unemp	0·002	0·000	−0·017	0·019	0·002	−0·001	0·001	0·000	0·005	−0·000	0·008	0·005	0·005	−0·009	−0·009	−0·022
	(0·5)	(0·0)	(1·1)	(1·8)	(0·3)	(0·1)	(0·1)	(0·0)	(0·4)	(0·0)	(2·4)	(0·5)	(0·5)	(0·8)	(0·7)	(2·4)
$\ln(N_{-1})$	−0·112	−0·187	−0·381	0·090	0·069	−0·067	0·020	−0·288	−0·168	−0·081	−0·011	−0·193	−0·139	−0·325	−0·280	−0·031
	(2·8)	(1·9)	(3·3)	(1·6)	(1·3)	(0·8)	(0·3)	(3·6)	(2·4)	(1·8)	(0·3)	(2·4)	(2·1)	(4·2)	(4·5)	(0·8)
LnPop	0·083	0·19	0·372	−0·122	−0·085	0·057	−0·043	0·256	0·160	0·082	0·011	0·166	0·099	0·369	0·255	0·030
	(2·1)	(2·0)	(3·0)	(2·1)	(1·6)	(0·6)	(0·7)	(3·3)	(2·1)	(1·8)	(0·3)	(2·1)	(1·4)	(5·0)	(4·0)	(0·7)
LnRinc	0·146	−0·442	−0·284	−0·159	−0·109	−0·292	0·043	−0·145	0·130	−0·035	0·135	−0·070	0·009	−0·165	0·254	0·009
	(1·8)	(1·7)	(1·0)	(0·8)	(0·7)	(1·3)	(0·3)	(1·1)	(0·6)	(0·3)	(2·1)	(0·4)	(0·0)	(0·8)	(1·0)	(0·1)
LnDens	−0·009	0·001	−0·004	0·027	0·016	−0·008	0·005	−0·011	−0·010	0·026	−0·022	0·010	−0·022	−0·020	−0·005	−0·002
	(1·3)	(0·0)	(0·2)	(1·6)	(1·1)	(0·4)	(0·3)	(1·0)	(0·6)	(2·5)	(3·8)	(0·7)	(1·0)	(1·2)	(0·2)	(0·1)
Poverty	0·007	−0·006	−0·011	−0·003	0·006	−0·003	0·000	−0·007	−0·001	−0·008	−0·003	−0·000	−0·009	−0·007	−0·002	0·006
	(3·0)	(0·8)	(1·2)	(0·5)	(1·1)	(0·5)	(0·0)	(1·9)	(0·2)	(2·0)	(1·6)	(0·0)	(1·6)	(1·3)	(0·2)	(1·0)
Urban	−0·001	−0·002	0·001	0·001	0·001	−0·001	0·001	−0·001	0·001	0·001	−0·001	0·002	0·002	−0·004	0·001	0·001
	(1·3)	(1·2)	(0·5)	(0·4)	(0·7)	(0·4)	(1·3)	(0·6)	(0·5)	(0·7)	(3·6)	(1·4)	(1·4)	(2·3)	(0·7)	(1·0)
Over65	−0·002	0·007	0·004	−0·011	0·004	−0·015	−0·005	−0·001	0·005	−0·009	−0·001	−0·001	−0·011	−0·008	−0·012	0·005
	(0·8)	(0·7)	(0·4)	(1·5)	(0·8)	(1·8)	(0·9)	(0·2)	(0·6)	(1·7)	(0·5)	(0·2)	(1·6)	(1·0)	(1·3)	(0·8)
R^2	0·557	0·454	0·401	0·361	0·361	0·523	0·431	0·562	0·324	0·502	0·702	0·339	0·483	0·627	0·585	0·214

Note: The dependent variable is the change in the logarithm of the number of establishments from 1993 to 1997. *T*-values between brackets. The state of Alaska and Washington DC are left out, leaving 49 observations for each regression.

CONCLUSION

The question of whether there is an important outflow from unemployment into self-employment and the start-up of new enterprises is of interest to labour economists, industrial economists and researchers in the field of regional development. In this paper we test for US states whether regions in which there is high unemployment show increases in the number of establishments in retail and consumer service industries in a subsequent four-year period. These industries are characterized by low barriers of entry and therefore provide an adequate test of the 'unemployment push' hypothesis. The general conclusion is that we fail to find an effect of unemployment after correcting for market opportunities. However, dependent upon the estimation procedure used, one or two industries (used merchandise stores and automotive repair shops) could show a positive effect of unemployment on (net) start-up activity. Many other cross-sectional regional studies have found a negative effect of unemployment on the start-up rate. Our results indicate that this result may be spurious as correcting for market opportunities may leave the effect insignificant. However, since we do not find an effect of unemployment on the change in the number of establishments across a broad range of low barrier industries, the 'unemployment push' hypothesis is left unconfirmed. KEEBLE and WALKER, 1994, come to the same conclusion using UK county data and correcting for some key determinants of new enterprise formation (but not industry structure). A qualification to the present study is that establishments without payroll (viz. having had no employees during any time of the year) are excluded in the data source. Further research using different data sources may shed light on the extent of bias as a result of omitting these smallest of firms.

EVANS and LEIGHTON, 1990, show evidence that unemployed workers are about twice as likely to start businesses as employed workers. Their result may be reconciled with the current evidence as: (1) they show that there is also a higher exit rate (within a couple of years) out of self-employment for unemployed workers; (2) unemployed people may start very small enterprises like home-based businesses that fail to grow and are not incorporated in the data. In either case we conclude that even when unemployed workers would be more inclined to become self-employed than employed workers, they are less successful in terms of growth and survival. Therefore, strong hopes that (regional) unemployment may partially be resolved by an outburst of new and successful ventures is unwarranted.[15] However, this study does suggest that entry in *very* low barrier industries may be an option for the unemployed. To corroborate this finding further research is required.

Acknowledgements – This paper was written while the author was visiting the Institute for Development Strategies, Indiana University, Bloomington, IN, as an Ameritech Research Scholar. The author is grateful to the Royal Netherlands Academy of Arts and Sciences (KNAW) for financial support. The author also thanks Marcus Dejardin and three referees for comments.

NOTES

1. AUDRETSCH and JIN, 1994, propose a reconciliation on the basis of economic theory. The study of EVANS and LEIGHTON, 1990, who show that unemployed workers are about twice as likely to start businesses as employed workers, has been considered important support for the 'unemployment push' hypothesis.
2. Entry of unemployed workers into the manufacturing sector is usually difficult. See, for example, VIVARELLI, 1994, who finds that, in a sample of mainly manufacturing start-ups in Northern Italy, less than 1% was previously unemployed. Although entry barriers in retailing and consumer services are relatively low, they are not absent. GABLE *et al.*, 1995, suggest that factors like availability of store locations and availability of qualified personnel are important barriers for new entrants.
3. CARREE and THURIK, 1999, develop the concept of 'carrying capacity' for the retail industry somewhat further.
4. Cross-border effects are not taken into account. Some states may benefit from densely populated areas in other states close to their borders. An example is Washington state benefiting from Vancouver being close to the border. The cross-border effects can generally be considered as zero-sum.
5. In fact, it is not unlikely that the error terms $\varepsilon_{i_1 j}$ and $\varepsilon_{i_2 j}$ for two industries i_1 and i_2 are correlated (either positively or negatively) across regions. However, we doubt whether a potential entrepreneur when comparing 'his/her' region to the national average will take other industries into account as well. Ordinary least squares is consistent even when the error terms are correlated across states.
6. For very specialized stores that only have a small percentage of people interested in their products a relatively high degree of urbanization or population density may be necessary to make the venture profitable. For such stores the variables will have a positive effect on the number of establishments.
7. We use a four-year period to allow for unemployed workers to need time to set up a new venture or to grow from a small home-based business into an establishment. The average correlation of the unemployment rate in 1993 with the relative change in the number of establishments between 1993 and 1994 equals $-0{\cdot}058$. The same correlation but for the period between 1993 and 1995, 1993 and 1996 and 1993 and 1997 (the period we actually use) is $-0{\cdot}082$, $-0{\cdot}039$ and $-0{\cdot}025$, respectively. From this perspective, the four-year period is the more likely one to give empirical support for the 'unemployment push' hypothesis.
8. WICKER and KING, 1989, show that about half of a sample of owners of new retail/service establishments in Southern California had been previously employed in

the same line of business. It indicates that some experience in the sector can be important to successfully start-up.

9. The author is grateful to a referee for suggesting this procedure.
10. CARREE and THURIK, 1994, provide a study of the effect of the nation-wide level of unemployment on the net entry rate in Dutch retailing industries. They find a significantly positive, but small and not very robust, effect.
11. The share of establishments run by firms with 500 or more employees in the total number of establishments ranges from 49·1% for motion picture theatres, 41·7% for optical goods stores and 34·9% for book stores to less than 1% for motorcycle dealers, florists and dance studios, schools and halls.
12. It may be the case that its product is too 'specialized' to be supplied in rural areas – see note 6.
13. The correlation between the first step error term and the unemployment rate is low for most industries and about equally distributed between positive and negative correlations. Out of 16 industries, there are only five that have a (rounded) correlation coefficient of $-0{\cdot}2$ or less or $+0{\cdot}2$ or higher. One industry (SIC 554) has a (rounded) correlation coefficient of $-0{\cdot}4$, two have a coefficient of $+0{\cdot}3$ (SIC 5941 and 5942), one has a coefficient of $+0{\cdot}2$ (SIC 784) and one has a coefficient of $-0{\cdot}2$ (SIC 791). Therefore, extent of 'market room' and rate of unemployment appear largely unrelated.
14. For one industry (dance studios, schools and halls, SIC 791), a negative effect of the unemployment rate is found, which is not significant at the 5% level, but which is significant at the 10% level. An investigation of regional variation in the effect of unemployment was pursued by dividing the states into 23 states in the South and West and 26 states in the Northeast and North-Central parts. The states belonging to the latter group are chosen to be those that are North of North Carolina, Tennessee, Arkansas and Oklahoma and East of Montana, Wyoming and Colorado. RIGBY and ESSLETZBICHLER, 2000, show that there is a difference between the states of the South and West in rates of plant entry and exit compared to the states in the 'snowbelt'. This may suggest that the barriers to entry and exit and, hence, possibly the effect of unemployment differ across the two regions. We divided the unemployment effect into one for the South/West states and one for the Northeast and North-Central states. However, in none of the 16 industries did we find a significantly different effect for the two regions.
15. It should be stressed that we have used data from a highly developed country with a relatively low unemployment rate (6·3% on average in 1993). Evidence for unemployed workers to be more likely to start new firms or establishments is much more likely to be found for lesser developed countries (LDCs) with high unemployment rates. Employees of large and medium sized firms and governmental agencies in such countries will be very reluctant to give up their positions to start new ventures. The vast majority of the (informal) ventures will be set up by workers who cannot obtain employment in the formal sector.

REFERENCES

AUDRETSCH D. B. and FRITSCH M. (1994) The geography of firm births in Germany, *Reg. Studies* **28**, 359–65.

AUDRETSCH D. B. and FRITSCH M. (1999) The industry component of regional new firm formation processes, *Rev. Ind. Organiz.* **15**, 239–52.

AUDRETSCH D. and JIN J. (1994) A reconciliation of the unemployment – new firm startup paradox, *Small Business Econ.* **6**, 381–85.

CARREE M. and THURIK R. (1994) The dynamics of entry, exit and profitability: an error correction approach for the retail industry, *Small Business Econ.* **6**, 107–16.

CARREE M. A. and THURIK A. R. (1999) The carrying capacity and entry and exit flows in retailing, *Int. J. Ind. Organiz.* **17**, 985–1,007.

EVANS D. S. and LEIGHTON L. S. (1990) Small business formation by unemployed and employed workers, *Small Business Econ.* **2**, 319–30.

FOTI A. and VIVARELLI M. (1994) An econometric test of the self-employment model: the case of Italy, *Small Business Econ.* **6**, 81–93.

GABLE M., TOPOL M. T., MATHIS S. and FISHER M. E. (1995) Entry barriers in retailing, *J. Retailing & Consumer Services* **2**, 211–21.

KEEBLE D. and WALKER S. (1994) New firms, small firms and dead firms: spatial patterns and determinants in the United Kingdom, *Reg. Studies* **28**, 411–27.

REYNOLDS P. D., MILLER B. and MAKI W. R. (1995) Explaining regional variation in business births and deaths: US 1976–88, *Small Business Econ.* **7**, 389–407.

REYNOLDS P. D., STOREY D. J. and WESTHEAD P. (1994) Cross-national comparisons of the variation in new firm formation rates: an editorial overview, *Reg. Studies* **28**, 343–46.

RIGBY D. L. and ESSLETZBICHLER J. (2000) Impacts of industry mix, technological change, selection and plant entry/exit on regional productivity growth, *Reg. Studies* **34**, 333–42.

STOREY D. J. (1991) The birth of new firms – does unemployment matter? A review of the evidence, *Small Business Econ.* **3**, 167–78.

VIVARELLI M. (1991) The birth of new enterprises, *Small Business Econ.* **3**, 215–23.

WICKER A. W. and KING J. C. (1989) Employment, ownership, and survival in microbusiness: a study of new retail and service establishments, *Small Business Econ.* **1**, 137–52.